A COMMON MISCONCEPTION CONCERNING THE "GREATER MEAN SQUARE RULE"

BU-38-M D.S. Robson April, 1953

Suppose $X_1, X_2, \ldots, X_m$ are normally and independently distributed chance variables with mean zero and unknown variance σ_x^2, and $Y_1, Y_2, \ldots, Y_n$ are normally and independently distributed with mean zero and unknown variance σ_y^2 and with $Y_1, Y_2, \ldots, Y_n$ independent of $X_1, X_2, \ldots, X_m$. We wish to test the hypothesis H_0 that $\sigma_x^2 = \sigma_y^2$ against the alternative H_1 that $\sigma_x^2 \neq \sigma_y^2$. The conventional test is the two-tail F-test, ϕ_1, of size 2ϵ, where

$$\phi_1(x_1,\ldots,x_m,y_1,\ldots,y_n|m,n,\epsilon) = \begin{cases} 0 \text{ (i.e., accept } H_0) \text{ if } F^{(m)}_{(n),\epsilon} < \dfrac{s_x^2}{s_y^2} < F^{(m)}_{(n),1-\epsilon} \\ 1 \text{ (i.e., reject } H_0) \text{ otherwise} \end{cases}$$

where $0 < \epsilon < 1/2$. If we attempt to carry out this test in practice, we face the difficulty that $F^{(m)}_{(n),\epsilon}$ has not been tabulated; we note, however, that the event $\frac{s_x^2}{s_y^2} < F^{(m)}_{(n),\epsilon}$ and the event $\frac{s_y^2}{s_x^2} > F^{(n)}_{(m),1-\epsilon}$ are identical (one and the same event) since

$$\epsilon = P\left[\frac{s_x^2}{s_y^2} < F^{(m)}_{(n),\epsilon}\right] = P\left[\frac{s_y^2}{s_x^2} > \frac{1}{F^{(m)}_{(n),\epsilon}}\right] = P\left[\frac{s_y^2}{s_x^2} > F^{(n)}_{(m),1-\epsilon}\right].$$

In practice, therefore, we use instead of ϕ_1 an equivalent test ϕ_2 where

$$\phi_2(x_1,\ldots,x_m,y_1,\ldots,y_n|m,n,\epsilon) = \begin{cases} 0 \text{ if } \dfrac{s_x^2}{s_y^2} < F^{(m)}_{(n),1-\epsilon} \text{ and } \dfrac{s_y^2}{s_x^2} < F^{(n)}_{(m),1-\epsilon}. \\ 1 \text{ otherwise} \end{cases}$$

Note that the event $\frac{s_x^2}{s_y^2} > F^{(m)}_{(n),1-\epsilon}$ and the event $\frac{s_y^2}{s_x^2} > F^{(n)}_{(m),1-\epsilon}$ are mutually exclusive since the latter is identical to $\frac{s_x^2}{s_y^2} < F^{(m)}_{(n),\epsilon}$, so that $\phi_2 = \phi_1$ identically in m, n, and ϵ.

Now consider the "greater mean square rule", ϕ_3, where

$$\phi_3(x_1, \ldots, x_m, y_1, \ldots, y_n | m, n, \epsilon) = \begin{cases} 0 \text{ if } \max\left(\frac{s_x^2}{s_y^2}, \frac{s_y^2}{s_x^2}\right) < \text{corresponding } F^{(\ )}_{(\ ), 1-\epsilon} . \\ 1 \text{ otherwise} \end{cases}$$

A common misconception concerning this test is that $\phi_3 = \phi_2$ identically in all arguments. We shall show that this is not an identity except over a certain range of ϵ — — in fact, it holds for all values of ϵ normally encountered in F-tables; e.g., $\epsilon \leq .10$.

Proof: Suppose, in fact, that the misconception is true, then it follows that

$$F^{(m)}_{(n), 1-\epsilon} > 1 \text{ for all } m, n, \text{ and } \epsilon, \; 0 < \epsilon < 1/2$$

since if there did exist a (m_0, n_0, ϵ_0) such that $F^{(m_0)}_{(n_0), 1-\epsilon_0} < 1$

then there would be a positive probability of the event

$$1 > \frac{s_x^2}{s_y^2} > F^{(m_0)}_{(n_0), 1-\epsilon_0}$$

and if this event occurred we would have $\phi_1 = \phi_2 = 1 \neq \phi_3 = 0$.

The proof is then completed when we actually specify (m_0, n_0, ϵ_0) such that $F^{(m_0)}_{(n_0), 1-\epsilon_0} < 1$; take, for example, $m_0 = 1$, $n_0 = 25$, $\epsilon_0 = .40$, in which case $F^{(1)}_{(25), .60} = .73 < 1$, and we have attained a contradiction.

The proof may be clarified by a graphic illustration of the above example where $m = 1$, $n = 25$, $\epsilon = .40$. Figure 1 is a sketch of the H_0 distribution of $\frac{s_x^2}{s_y^2} = F^{(1)}_{(25)}$; the shaded areas comprise the rejection region of the test ϕ_1. Figure 2 sketches the H_0 distribution of $\frac{s_y^2}{s_x^2} = F^{(25)}_{(1)}$ with the corresponding shaded rejection region; thus, the upper shaded area of Figure 1 and the upper shaded area of Figure 2 together comprise the rejection region of the test ϕ_2. Now, suppose that in a particular experiment we observe a value of $\frac{s_x^2}{s_y^2}$ somewhere between .73 and 1, say $\frac{s_x^2}{s_y^2} = .95$, or

$\frac{s_y^2}{s_x^2} = \frac{1}{.95} = 1.1$. Since .95 lies in the upper shaded area of Figure 1 then both ϕ_1 and ϕ_2 would reject H_o; ϕ_3, however, would compare $\max\left(\frac{s_x^2}{s_y^2}, \frac{s_y^2}{s_x^2}\right) = \max(.95, 1.1) = 1.1$ with the corresponding F value 3.5, and accept H_o since $1.1 < 3.5$.

Figure 1. Schematic representation of the H_o distribution of $\frac{s_x^2}{s_y^2} = F_{(25)}^{(1)}$

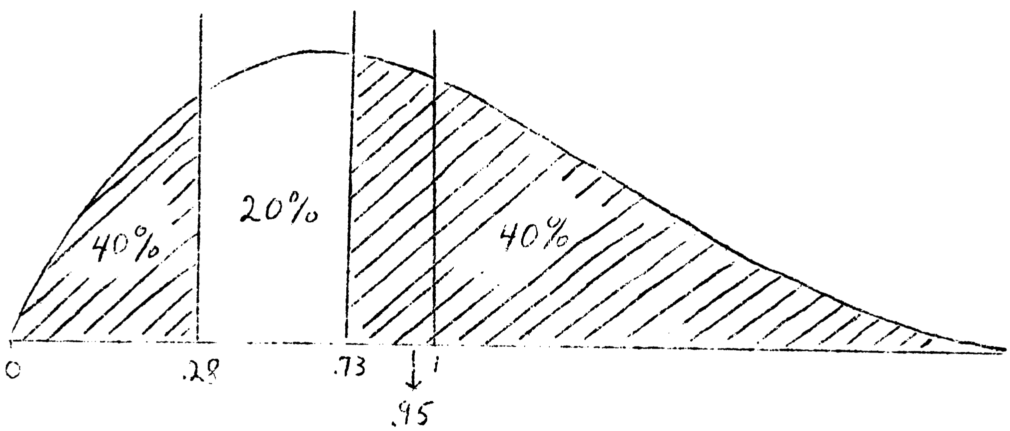
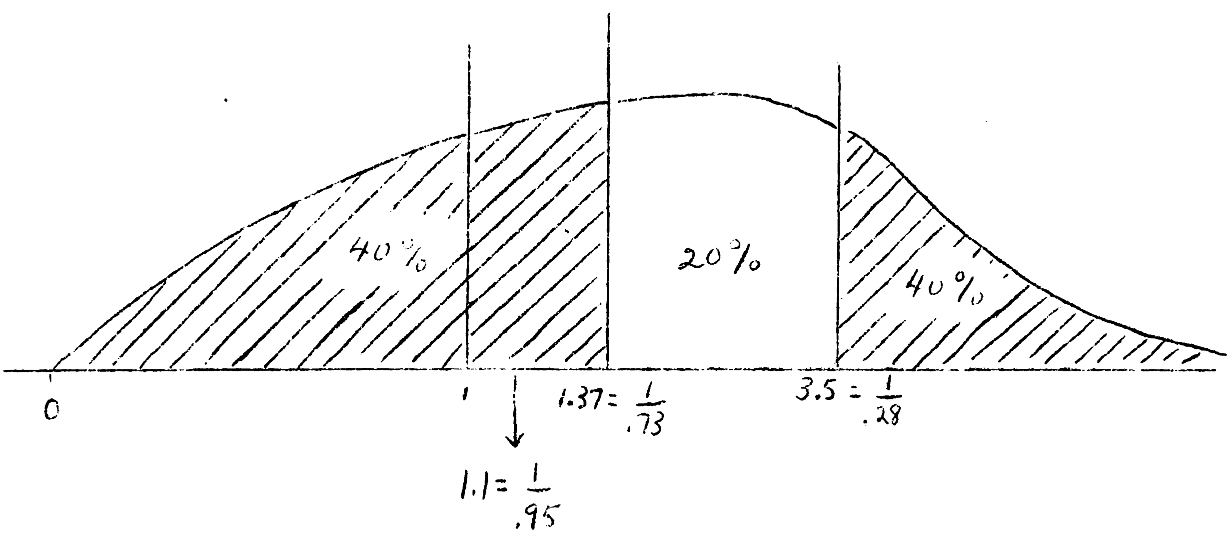


Figure 2. Schematic representation of the H_o distribution of $\frac{s_y^2}{s_x^2} = F_{(1)}^{(25)}$



We observe from the above proof that for a fixed ϵ_o,

$$\phi_2(x_1,\ldots, x_m, y_1,\ldots, y_n|m,n,\epsilon_o) \equiv \phi_3(x_1,\ldots, x_m, y_1,\ldots, y_n|m,n,\epsilon_o)$$

implies that

$$F^{(m)}_{(n),1-\epsilon_o} > 1 \text{ for all } m, n.$$

We now note that the opposite implication also holds; i.e., for fixed ϵ_o

$$F^{(m)}_{(n),1-\epsilon_o} > 1 \text{ for all } m, n$$

implies that

$$\phi_2(x_1,\ldots, x_m, y_1,\ldots, y_n|m,n,\epsilon_o) \equiv \phi_3(x_1,\ldots, x_m, y_1,\ldots, y_n|m,n,\epsilon_o).$$

This latter implication follows from the fact that

$$\min\left(\frac{s_x^2}{s_y^2}, \frac{s_y^2}{s_x^2}\right) < 1 < \max\left(\frac{s_x^2}{s_y^2}, \frac{s_y^2}{s_x^2}\right);$$

thus, if one of the ratios $\frac{s_x^2}{s_y^2}$ and $\frac{s_y^2}{s_x^2}$ is to exceed the corresponding $F^{(\;)}_{(\;),1-\epsilon_o}$ then it must be the ratio which is larger than 1 since both $F^{(m)}_{(n),1-\epsilon}$ and $F^{(n)}_{(m),1-\epsilon_o}$ are larger than 1. Hence, in using the test ϕ_2 under these conditions it would be sufficient to consider only the larger of the two ratios, and this is precisely the test ϕ_3.

We have thus shown that if ϵ is such that $F^{(m)}_{(n),1-\epsilon} > 1$ for all m, n then for any pair of sample sizes m and n the three tests ϕ_1, ϕ_2, ϕ_3 are identical. Reference to the tabulated F distribution shows, in particular, that all $\epsilon \leq .10$ satisfy this condition. Since the size (2ϵ) of the test rarely exceeds $2(.10) = .20$ in practice, a search for any larger values of ϵ $(\epsilon > .10)$ which satisfy this condition would be of little value.

In summary, we have shown that the "greater mean square rule" <u>does</u> provide an exact test of size 2ϵ <u>for all values of ϵ normally encountered in practice</u> (e.g., $\epsilon \leq .10$) but that those statisticians who thought the test good for <u>any</u> value of ϵ $(0 < \epsilon < 1/2)$ harbored a misconcept.